

RECEIVED

MAR 07 2002

TECH CENTER 1600/2900

#12/1008 1636
3/8/02

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/810,836A

TIME: 11:15:46

Input Set : A:\SEQLIST514442001200.TXT

Output Set: N:\CRF3\02212002\I810836A.raw

ENTERED

```

4 <110> APPLICANT: Broun, Pierre
6 <120> TITLE OF INVENTION: METHOD FOR MODIFYING A BIOSYNTHETIC
7   PATHWAY
9 <130> FILE REFERENCE: 514442001200/MBI0032
11 <140> CURRENT APPLICATION NUMBER: US 09/810,836A
12 <141> CURRENT FILING DATE: 2001-03-16
14 <160> NUMBER OF SEQ ID NOS: 14
16 <170> SOFTWARE: FastSEQ for Windows Version 4.0
18 <210> SEQ ID NO: 1
19 <211> LENGTH: 1239
20 <212> TYPE: DNA
21 <213> ORGANISM: Arabidopsis thaliana
23 <220> FEATURE:
24 <221> NAME/KEY: CDS
25 <222> LOCATION: (6)...(1091)
26 <223> OTHER INFORMATION: G993
28 <400> SEQUENCE: 1
29 caaat atg gaa tac agc tgt gta gac gac agt agt aca acg tca gaa tct      50
30   Met Glu Tyr Ser Cys Val Asp Asp Ser Ser Thr Thr Ser Glu Ser
31       1           5           10           15
33 ctc tcc atc tct act act cca aag ccg aca acg acg acg gag aag aaa      98
34 Leu Ser Ile Ser Thr Thr Pro Lys Pro Thr Thr Thr Thr Glu Lys Lys
35           20           25           30
37 ctc tct tct ccg ccg gcg acg tcg atg cgt ctc tac aga atg gga agc      146
38 Leu Ser Ser Pro Pro Ala Thr Ser Met Arg Leu Tyr Arg Met Gly Ser
39           35           40           45
41 ggc gga agc agc gtc gtt ttg gat tca gag aac ggc gtc gag acc gag      194
42 Gly Gly Ser Ser Val Val Leu Asp Ser Glu Asn Gly Val Glu Thr Glu
43       50           55           60
45 tca cgt aag ctt cct tcg tcg aaa tat aaa ggc gtt gtg cct cag cct      242
46 Ser Arg Lys Leu Pro Ser Ser Lys Tyr Lys Gly Val Val Pro Gln Pro
47       65           70           75
49 aac gga aga tgg gga gct cag att tac gag aag cat cag cga gtt tgg      290
50 Asn Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys His Gln Arg Val Trp
51   80           85           90           95
53 ctc ggt act ttc aac gag gaa gaa gaa gct gcg tct tct tac gac atc      338
54 Leu Gly Thr Phe Asn Glu Glu Glu Glu Ala Ala Ser Ser Tyr Asp Ile
55       100           105           110
57 gcc gtg agg aga ttc cgc ggc cgc gac gcc gtc act aac ttc aaa tct      386
58 Ala Val Arg Arg Phe Arg Gly Arg Asp Ala Val Thr Asn Phe Lys Ser
59       115           120           125
61 caa gtt gat gga aac gac gcc gaa tcg gct ttt ctt gac gct cat tct      434
62 Gln Val Asp Gly Asn Asp Ala Glu Ser Ala Phe Leu Asp Ala His Ser

```

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/810,836A

TIME: 11:15:46

Input Set : A:\SEQLIST514442001200.TXT

Output Set: N:\CRF3\02212002\I810836A.raw

```

63          130          135          140
65 aaa gct gag atc gtg gat atg ttg agg aaa cac act tac gcc gat gag      482
66 Lys Ala Glu Ile Val Asp Met Leu Arg Lys His Thr Tyr Ala Asp Glu
67          145          150          155
69 ttt gag cag agt aga cgg aag ttt gtt aac ggc gac gga aaa cgc tct      530
70 Phe Glu Gln Ser Arg Arg Lys Phe Val Asn Gly Asp Gly Lys Arg Ser
71 160          165          170          175
73 ggg ttg gag acg gcg acg tac gga aac gac gct gtt ttg aga gcg cgt      578
74 Gly Leu Glu Thr Ala Thr Tyr Gly Asn Asp Ala Val Leu Arg Ala Arg
75          180          185          190
77 gag gtt ttg ttc gag aag act gtt acg ccg agc gac gtc ggg aag ctg      626
78 Glu Val Leu Phe Glu Lys Thr Val Thr Pro Ser Asp Val Gly Lys Leu
79          195          200          205
81 aac cgt tta gtg ata ccg aaa caa cac gcg gag aag cat ttt ccg tta      674
82 Asn Arg Leu Val Ile Pro Lys Gln His Ala Glu Lys His Phe Pro Leu
83          210          215          220
85 ccg gcg atg acg acg gcg atg ggg atg aat ccg tct ccg acg aaa ggc      722
86 Pro Ala Met Thr Thr Ala Met Gly Met Asn Pro Ser Pro Thr Lys Gly
87          225          230          235
89 gtt ttg att aac ttg gaa gat aga aca ggg aaa gtg tgg cgg ttc cgt      770
90 Val Leu Ile Asn Leu Glu Asp Arg Thr Gly Lys Val Trp Arg Phe Arg
91 240          245          250          255
93 tac agt tac tgg aac agc agt caa agt tac gtg ttg acc aag ggc tgg      818
94 Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly Trp
95          260          265          270
97 agc cgg ttc gtt aaa gag aag aat ctt cga gcc ggt gat gtg gtt tgt      866
98 Ser Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val Val Cys
99          275          280          285
101 ttc gag aga tca acc gga cca gac cgg caa ttg tat atc cac tgg aaa      914
102 Phe Glu Arg Ser Thr Gly Pro Asp Arg Gln Leu Tyr Ile His Trp Lys
103          290          295          300
105 gtc cgg tct agt ccg gtt cag act gtg gtt agg cta ttc gga gtc aac      962
106 Val Arg Ser Ser Pro Val Gln Thr Val Val Arg Leu Phe Gly Val Asn
107          305          310          315
109 att ttc aat gtg agt aac gag aaa cca aac gac gtc gca gta gag tgt      1010
110 Ile Phe Asn Val Ser Asn Glu Lys Pro Asn Asp Val Ala Val Glu Cys
111 320          325          330          335
113 gtt ggc aag aag aga tct cgg gaa gat gat ttg ttt tcg tta ggg tgt      1058
114 Val Gly Lys Lys Arg Ser Arg Glu Asp Asp Leu Phe Ser Leu Gly Cys
115          340          345          350
117 tcc aag aag cag gcg att atc aac atc ttg tga caaattcttt ttttttggtt      1111
118 Ser Lys Lys Gln Ala Ile Ile Asn Ile Leu *
119          355          360
121 tttttcttca atttgtttct cttttttcaa tattttgtat tgaaatgaca agttgtaaat      1171
122 taggacaaga caagaaaaaa tgacaactag acaaaatagt ttttgtttaa aaaaaaaaaa      1231
123 aaaaaaaaaa      1239
125 <210> SEQ ID NO: 2
126 <211> LENGTH: 361
127 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/810,836A

TIME: 11:15:46

Input Set : A:\SEQLIST514442001200.TXT

Output Set: N:\CRF3\02212002\I810836A.raw

128 <213> ORGANISM: Arabidopsis thaliana

130 <400> SEQUENCE: 2

```

131 Met Glu Tyr Ser Cys Val Asp Asp Ser Ser Thr Thr Ser Glu Ser Leu
132 1 5 10 15
133 Ser Ile Ser Thr Thr Pro Lys Pro Thr Thr Thr Thr Glu Lys Lys Leu
134 20 25 30
135 Ser Ser Pro Pro Ala Thr Ser Met Arg Leu Tyr Arg Met Gly Ser Gly
136 35 40 45
137 Gly Ser Ser Val Val Leu Asp Ser Glu Asn Gly Val Glu Thr Glu Ser
138 50 55 60
139 Arg Lys Leu Pro Ser Ser Lys Tyr Lys Gly Val Val Pro Gln Pro Asn
140 65 70 75 80
141 Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys His Gln Arg Val Trp Leu
142 85 90 95
143 Gly Thr Phe Asn Glu Glu Glu Glu Ala Ala Ser Ser Tyr Asp Ile Ala
144 100 105 110
145 Val Arg Arg Phe Arg Gly Arg Asp Ala Val Thr Asn Phe Lys Ser Gln
146 115 120 125
147 Val Asp Gly Asn Asp Ala Glu Ser Ala Phe Leu Asp Ala His Ser Lys
148 130 135 140
149 Ala Glu Ile Val Asp Met Leu Arg Lys His Thr Tyr Ala Asp Glu Phe
150 145 150 155 160
151 Glu Gln Ser Arg Arg Lys Phe Val Asn Gly Asp Gly Lys Arg Ser Gly
152 165 170 175
153 Leu Glu Thr Ala Thr Tyr Gly Asn Asp Ala Val Leu Arg Ala Arg Glu
154 180 185 190
155 Val Leu Phe Glu Lys Thr Val Thr Pro Ser Asp Val Gly Lys Leu Asn
156 195 200 205
157 Arg Leu Val Ile Pro Lys Gln His Ala Glu Lys His Phe Pro Leu Pro
158 210 215 220
159 Ala Met Thr Thr Ala Met Gly Met Asn Pro Ser Pro Thr Lys Gly Val
160 225 230 235 240
161 Leu Ile Asn Leu Glu Asp Arg Thr Gly Lys Val Trp Arg Phe Arg Tyr
162 245 250 255
163 Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly Trp Ser
164 260 265 270
165 Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val Val Cys Phe
166 275 280 285
167 Glu Arg Ser Thr Gly Pro Asp Arg Gln Leu Tyr Ile His Trp Lys Val
168 290 295 300
169 Arg Ser Ser Pro Val Gln Thr Val Val Arg Leu Phe Gly Val Asn Ile
170 305 310 315 320
171 Phe Asn Val Ser Asn Glu Lys Pro Asn Asp Val Ala Val Glu Cys Val
172 325 330 335
173 Gly Lys Lys Arg Ser Arg Glu Asp Asp Leu Phe Ser Leu Gly Cys Ser
174 340 345 350
175 Lys Lys Gln Ala Ile Ile Asn Ile Leu
176 355 360
178 <210> SEQ ID NO: 3

```

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/810,836A

TIME: 11:15:46

Input Set : A:\SEQLIST514442001200.TXT

Output Set: N:\CRF3\02212002\I810836A.raw

```

179 <211> LENGTH: 1226
180 <212> TYPE: DNA
181 <213> ORGANISM: Arabidopsis thaliana
183 <220> FEATURE:
184 <221> NAME/KEY: CDS
185 <222> LOCATION: (111)...(989)
186 <223> OTHER INFORMATION: G1845
188 <400> SEQUENCE: 3
189 aagacataat tttctctgtt ttcctagctc tctcctctca aattcttcca ttgctctctg      60
190 ttttggcaaa tcgtgaactg ccacgtcttt aaggcatcag tgaagcaaag atg gac      116
191                                     Met Asp
192                                     1
194 ttt gac gag gag cta aat ctt tgt att acg aaa ggt aaa aat gtt gat      164
195 Phe Asp Glu Glu Leu Asn Leu Cys Ile Thr Lys Gly Lys Asn Val Asp
196         5                10                15
198 cat tct ttt gga gga gaa gct tct tcc acg tcc cca aga tct atg aag      212
199 His Ser Phe Gly Gly Glu Ala Ser Ser Thr Ser Pro Arg Ser Met Lys
200        20                25                30
202 aaa atg aag agt cct agt cgt cct aaa ccc tat ttc caa tcc tct tct      260
203 Lys Met Lys Ser Pro Ser Arg Pro Lys Pro Tyr Phe Gln Ser Ser Ser
204       35                40                45                50
206 tct cct tat tcg tta gag gct ttc cct ttt tct ctc gat cca aca ctt      308
207 Ser Pro Tyr Ser Leu Glu Ala Phe Pro Phe Ser Leu Asp Pro Thr Leu
208             55                60                65
210 cag aat cag caa caa caa ctc gga tca tac gtt ccg gta ctt gag caa      356
211 Gln Asn Gln Gln Gln Gln Leu Gly Ser Tyr Val Pro Val Leu Glu Gln
212             70                75                80
214 cga caa gac ccg aca atg caa ggc cag aag caa atg atc tcc ttt agt      404
215 Arg Gln Asp Pro Thr Met Gln Gly Gln Lys Gln Met Ile Ser Phe Ser
216             85                90                95
218 cct caa caa caa caa cag cag cag cag tat atg gcc cag tac tgg agt      452
219 Pro Gln Gln Gln Gln Gln Gln Gln Gln Tyr Met Ala Gln Tyr Trp Ser
220        100                105                110
222 gac aca ttg aat ctg agt cca aga gga aga atg atg atg atg atg agc      500
223 Asp Thr Leu Asn Leu Ser Pro Arg Gly Arg Met Met Met Met Met Ser
224       115                120                125                130
226 caa gaa gct gtt caa cct tac atc gca acg aag ctg tac aga gga gtg      548
227 Gln Glu Ala Val Gln Pro Tyr Ile Ala Thr Lys Leu Tyr Arg Gly Val
228             135                140                145
230 aga caa cgt caa tgg gga aaa tgg gtc gca gag atc cgt aag cca cga      596
231 Arg Gln Arg Gln Trp Gly Lys Trp Val Ala Glu Ile Arg Lys Pro Arg
232             150                155                160
234 agc agg gca cgt ctt tgg ctt ggt acc ttt gat aca gct gaa gaa gct      644
235 Ser Arg Ala Arg Leu Trp Leu Gly Thr Phe Asp Thr Ala Glu Glu Ala
236             165                170                175
238 gcc atg gcc tac gac cgc caa gcc ttc aaa tta cga ggc cac agc gca      692
239 Ala Met Ala Tyr Asp Arg Gln Ala Phe Lys Leu Arg Gly His Ser Ala
240             180                185                190
242 aca ctg aat ttc ccg gag cat ttt gtg aat aag gaa agc gag ctg cat      740

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/810,836A

DATE: 02/21/2002

TIME: 11:15:46

Input Set : A:\SEQLIST514442001200.TXT

Output Set: N:\CRF3\02212002\I810836A.raw

```

243 Thr Leu Asn Phe Pro Glu His Phe Val Asn Lys Glu Ser Glu Leu His
244 195                200                205                210
246 gat tca aac tcg tcg gat cag aaa gaa cct gaa acg cca cag cca agc      788
247 Asp Ser Asn Ser Ser Asp Gln Lys Glu Pro Glu Thr Pro Gln Pro Ser
248                215                220                225
250 gag gtt aac ttg gag agc aag gaa cta ccg gtg att gat gtt ggg aga      836
251 Glu Val Asn Leu Glu Ser Lys Glu Leu Pro Val Ile Asp Val Gly Arg
252                230                235                240
254 gag gaa ggt atg gct gag gca tgg tac aat gcc att aca tcg gga tgg      884
255 Glu Glu Gly Met Ala Glu Ala Trp Tyr Asn Ala Ile Thr Ser Gly Trp
256                245                250                255
258 ggt cct gaa agt cct ctt tgg gat gat ttg gat agt tct cat cag ttt      932
259 Gly Pro Glu Ser Pro Leu Trp Asp Asp Leu Asp Ser Ser His Gln Phe
260                260                265                270
262 tca tca gaa agc tca tct tct tct cct ctc tct tgt cct atg agg cct      980
263 Ser Ser Glu Ser Ser Ser Ser Ser Pro Leu Ser Cys Pro Met Arg Pro
264 275                280                285                290
266 ttc ttt tga aaaagtttat aaaccacat tgtgtttag gttatagttt      1029
267 Phe Phe *
270 agggttatgc tcattggcat ttggatggag gcaatttttg tgatctccca ttccaccaca      1089
271 tatcagtcac tatatgtgtc taccttttct ctgtatttct atcattatca ttgtttttat      1149
272 tatgtgtctg tatgtgttct cctattgcta catacataga tgtcctcttt gttcaaaaaa      1209
273 aaaaaaaaaa aaaaaaa
275 <210> SEQ ID NO: 4
276 <211> LENGTH: 292
277 <212> TYPE: PRT
278 <213> ORGANISM: Arabidopsis thaliana
280 <400> SEQUENCE: 4
281 Met Asp Phe Asp Glu Glu Leu Asn Leu Cys Ile Thr Lys Gly Lys Asn
282 1          5          10          15
283 Val Asp His Ser Phe Gly Gly Glu Ala Ser Ser Thr Ser Pro Arg Ser
284                20          25          30
285 Met Lys Lys Met Lys Ser Pro Ser Arg Pro Lys Pro Tyr Phe Gln Ser
286                35          40          45
287 Ser Ser Ser Pro Tyr Ser Leu Glu Ala Phe Pro Phe Ser Leu Asp Pro
288                50          55          60
289 Thr Leu Gln Asn Gln Gln Gln Gln Leu Gly Ser Tyr Val Pro Val Leu
290 65          70          75          80
291 Glu Gln Arg Gln Asp Pro Thr Met Gln Gly Gln Lys Gln Met Ile Ser
292                85          90          95
293 Phe Ser Pro Gln Gln Gln Gln Gln Gln Gln Tyr Met Ala Gln Tyr
294                100         105         110
295 Trp Ser Asp Thr Leu Asn Leu Ser Pro Arg Gly Arg Met Met Met Met
296                115         120         125
297 Met Ser Gln Glu Ala Val Gln Pro Tyr Ile Ala Thr Lys Leu Tyr Arg
298                130         135         140
299 Gly Val Arg Gln Arg Gln Trp Gly Lys Trp Val Ala Glu Ile Arg Lys
300 145         150         155         160
301 Pro Arg Ser Arg Ala Arg Leu Trp Leu Gly Thr Phe Asp Thr Ala Glu

```

VERIFICATION SUMMARY

DATE: 02/21/2002

PATENT APPLICATION: US/09/810,836A

TIME: 11:15:47

Input Set : A:\SEQLIST514442001200.TXT

Output Set: N:\CRF3\02212002\I810836A.raw